

IN THE CLAIMS:

1. (Currently Amended) A storage system for use in a storage system cluster, the storage system comprising:

a first server and a second server, wherein the second server is a cluster partner to the first server; and

a storage operating system operating on the first server, the storage operating system including a cluster connection manager configured to create, destroy, and maintain one or more communication sessions with the cluster partner, the cluster connection manager operatively interconnected with a set of cluster connection manager clients, where each cluster connection manager client is a process executing on the storage system, and wherein the cluster connection manager is further configured to create, destroy, and maintain a virtual interface connections between a cluster connection manager client on the first server with a cluster connection manager client on the second server to form a peer process between the cluster connection manager clients, wherein the virtual interface connection allows remote direct memory access (RDMA) operations that allow the cluster connection manager operating on the first server to directly access memory regions of the cluster partner operating on the second server.

2. (Original) The storage system of claim 1 wherein one of the set of communication clients comprises a failover monitor.

3. (Original) The storage system of claim 1 wherein one of the set of cluster connection manager clients comprises a non-volatile random access memory shadowing process.

- 1 4. (Original) The storage system of claim 1 wherein the cluster connection manager is
2 further adapted to perform connection management operations in response to communi-
3 cations from the connection manager clients.
- 1 5. (Original) The storage system of claim 4 wherein the communications comprise an ap-
2 plication program interface function call.
- 1 6. (Original) The storage system of claim 1 wherein the cluster connection manager is
2 further adapted to load balance the one or more communication sessions over a plurality
3 of cluster interconnect devices.
- 1 7. (Original) The storage system of claim 1 wherein the cluster connection manager is
2 further adapted to perform a failover procedure for one or more communication sessions
3 from a failed cluster interconnect device to an operational cluster interconnect device.
- 1 8. (Original) The storage system of claim 1 wherein the cluster connection manager is
2 operatively interconnected with a plurality of cluster interconnect devices.
- 1 9. (Original) The storage system of claim 1 wherein the storage operating system com-
2 prises a plurality of cluster connection managers.
- 1 10. (Currently Amended) A storage operating system, executing on a storage system, the
2 storage operating system comprising:
3 a cluster connection manager configured to manage a set of peer-to-peer connec-
4 tions associated with a set of cluster connection manager clients executing on the storage

5 system, wherein the cluster connection manger is further configured to create, destroy,
6 and maintain a virtual interface connection between a cluster connection manager client
7 on a first server with a cluster connection manager client on a second server to form a
8 peer process between cluster connection manager clients, wherein the virtual interface
9 connection allows remote direct memory access (RDMA) operations that allow the clus-
10 ter connection manager operating on the first server to directly access memory regions of
11 the cluster partner operating on the second server.

1 11. (Original) The storage operating system of claim 10 wherein the set of cluster con-
2 nection manager clients comprises a failover monitor.

1 12. (Original) The storage operating system of claim 10 wherein the cluster connection
2 manager is further adapted to perform load balancing of the set of peer-to-peer connec-
3 tions over a plurality of cluster interconnect devices.

1 13. (Original) The storage operating system of claim 10 wherein the cluster connection
2 manager is further adapted to failover the set of peer-to-peer connections from a failed
3 cluster interconnect device to an operational cluster interconnect device.

1 14. (Currently Amended) A method for initiating a peer-to-peer communication session,
2 comprising:

3 creating, using a cluster connection manager executing on a first server, an initial
4 connection with a cluster partner on a second server;
5 exchanging a set of peer connection information;
6 passing a set of cluster connection manager client information to the cluster part-
7 ner, wherein the set of cluster connection manager client information includes at

8 least one virtual interface and any memory requirements for each cluster manager
9 client;
10 creating a set of appropriate communication ports using the set of cluster connec-
11 tion manager client information, wherein the virtual interface connection allows remote
12 direct memory access (RDMA) operations that allow the cluster connection manager op-
13 erating on the first server to directly access memory regions of the cluster partner operat-
14 ing on the second server;
15 alerting the cluster partner of a ready status; and
16 alerting a set of cluster connection manager clients that the cluster partner is in a
17 ready state.

1 15. (Original) The method of claim 14 wherein the set of clients comprises a failover
2 monitor process.

1 16. (Original) The method of claim 14 wherein the set of peer connection information
2 comprises a version number.

1 17. (Previously Presented) The method of claim 14 wherein the step of passing a set of
2 client information to the cluster partner further comprises:

3 collecting, from a set of clients, the set of client information; and
4 transferring the collected set of client information to the cluster.

1 18. (Original) The method of claim 17 wherein the client information comprises a num-
2 ber of communication ports required.

1 19. (Original) The method of claim 17 wherein the set of client information further com-
2 prises an amount of memory requested by a particular client.

1 20. (Previously Presented) The method of claim 14 wherein the step of creating an initial
2 connection further comprises using remote direct memory access primitives to create the
3 initial connection.

1 21. (Previously Presented) The method of claim 14 wherein the step of creating an initial
2 connection further comprises performing a series of remote direct memory access opera-
3 tions to create the initial connection.

1 22. (Currently Amended) A method for terminating a peer-to-peer communication ses-
2 sion, comprising:

3 alerting, using a cluster connection manager executing on a storage system, a set
4 of clients of an impending termination of the communication session;

5 closing, by the clients, a set of communication ports associated with the commu-
6 nication session, wherein the set of communication ports comprise a set of virtual inter-
7 face connections; and

8 performing an initialization of a peer-to-peer communication session procedure,
9 wherein each virtual interface connection allows remote direct memory access (RDMA)
10 operations that allow the cluster connection manager operating on the first server to di-
11 rectly access memory regions of the cluster partner operating on the second server.

1 23. (Original) The method of claim 22 wherein the set of communication ports comprises
2 a set of virtual interface connections.

1 24. (Original) The method of claim 22 wherein the set of clients comprises a failover
2 monitor.

1 25. (Currently Amended) A storage system cluster, comprising:

2 a first storage system of the storage system cluster, the first storage system having
3 a first disk shelf and a first cluster connection manager to manage data flow from/to an
4 external source to the first disk shelf, wherein the first cluster connection manger is con-
5 figured to create, destroy, and maintain a virtual interface connections between a cluster
6 connection manager client on the first storage system with a cluster connection manager
7 client on a second storage system to form a peer process between the cluster connection
8 manager clients;

9 a second storage system of the storage system cluster, the second storage system
10 having a second disk shelf and a second cluster connection manager to manage data flow
11 from/to an external source to the second disk shelf; and

12 the first cluster connection manager to shift data flow from/to the first disk shelf
13 to the second disk shelf upon an event condition, wherein the virtual interface connection
14 allows remote direct memory access (RDMA) operations that allow the cluster connec-
15 tion manager operating on the first server to directly access memory regions of the cluster
16 partner operating on the second server.

1 26. (Previously Presented) The storage system cluster of claim 25 wherein the event con-
2 dition is a failed interconnect driver connected the first cluster connection manager.

1 27. (Previously Presented) The storage system cluster of claim 25 wherein the event con-
2 dition is a load-balancing condition.

1 28. (Currently Amended) A system configured to manage reliable peer communication
2 among storage systems in a clustered environment, the system comprising:

3 one or more peer processes executing on each storage system partner; and

4 a cluster connection manager executing on each storage system partner, the clus-
5 ter connection manager creating a set of peer-to-peer connections between the one or
6 more peer processes executing on each storage system, wherein the cluster connection
7 manager is provided to reliably create virtual interface connections between peer proc-
8 esses executing on the storage system partners over a cluster interconnect without requir-
9 ing a storage operating system executing on each storage system to be fully active or
10 functioning, wherein the virtual interface connection allows remote direct memory access
11 (RDMA) operations that allow the cluster connection manager operating on the first
12 server to directly access memory regions of the cluster partner operating on the second
13 server.

1 29. (Currently Amended) A computer readable medium for initiating a peer-to-peer
2 communication session, the computer readable medium including program instructions
3 executed by a processor for performing the steps of:

4 creating, using a cluster connection manager executing on a first server, an initial
5 connection with a cluster partner on a second server;

6 exchanging a set of peer connection information;

7 passing a set of cluster connection manager client information to the cluster part-
8 ner, wherein the set of cluster connection manager client information includes at
9 least one virtual interface and any memory requirements for each cluster manager
10 client;

11 creating a set of appropriate communication ports using the set of cluster connec-
12 tion manager client information, wherein the virtual interface connection allows remote

direct memory access (RDMA) operations that allow the cluster connection manager operating on the first server to directly access memory regions of the cluster partner operating on the second server;

alerting the cluster partner of a ready status; and

alerting a set of cluster connection manager clients that the cluster partner is in a ready state.

30. (Currently Amended) A computer readable medium for terminating a peer-to-peer communication session, the computer readable medium including program instructions executed by a processor for performing the steps of:

alerting, using a cluster connection manager executing on a storage system, a set of clients of an impending termination of the communication session;

closing, by the clients, a set of communication ports associated with the communication session, wherein the set of communication ports comprise a set of virtual interface connections; and

performing an initialization of a peer-to-peer communication session procedure, wherein the virtual interface connection allows remote direct memory access (RDMA) operations that allow the cluster connection manager operating on the first server to directly access memory regions of the cluster partner operating on the second server.

31. – 37. (Cancelled)

38. (Currently Amended) A storage operating system, executing on a storage system, the storage operating system comprising:

one or more peer processes executing on each storage system partner;

a plurality of cluster interconnect drivers executing on the storage system; and

6 ~~one or more a second~~ cluster connection managers configured to detect a high
7 bandwidth load on a first cluster connection manager through remote direct memory ac-
8 cess (RDMA) operations that allow the second cluster connection manager operating on a
9 second server to directly access memory regions of the first cluster connection manager
10 operating on a first server- and in response to detecting a high band width load, utilize a
11 second cluster connection manager to access each storage system partner.